# Multiplicative aggregation in managerial multi-attribute decision making

Maksim Goman* and Stefan Koch

Institute of Business Informatics – Information Engineering,
Johannes Kepler University,
Altenberger Street 69, 4040 Linz, Austria
Email: maksim.goman@jku.at
Email: stefan.koch@jku.at
*Corresponding author

**Abstract:** This paper focuses on aggregated performance of alternatives for management decision making. Assuming non-comparable criteria, we propose a composite indicator (CI) based on weighted product instead of commonly applied weighted average (WA). We extensively compare WA and CI in a real-world example of strategic decision making problem regarding enterprise resource planning system upgrade. The CI shows robustness to data scale change. User preference for a decision support method was examined based on complexity perception and willingness to use. The users are more likely to understand simple methods and apply them rather than methods that they do not comprehend, and the proposed approach is rated 'statistically not worse' than WA in this regard. Our findings should help managers in practical multi-attribute problems where alternative ranking based on a number of non-comparable properties is required. The alternative's rank is obtained in a mathematically correct way, and the aggregation does not need data normalisation.

**Keywords:** multi-attribute decision making; MADM; strategic decision making; software selection.

**Reference** to this paper should be made as follows: Goman, M. and Koch, S. (2021) 'Multiplicative aggregation in managerial multi-attribute decision making', *Int. J. Multicriteria Decision Making*, Vol. 8, No. 3, pp.233–255.

**Biographical notes:** Maksim Goman received his Diploma in Computer Science from the Belarussian State University of Informatics and Radioelectronics in 2001 and Master's degree from TU Ilmenau, in 2017. He started his academic career at the Academy of Sciences of Belarus in 2002 and has more than ten publications. He worked at engineering positions, as an IT auditor (CISA certification) and risk analyst at banks. Currently, he is a researcher at the Johannes Kepler University Linz, Austria.

Stefan Koch is a Professor and Chair at the Department of Business Informatics – Information Engineering, Johannes Kepler University Linz. He received his PhD from the WU – Vienna University of Economics and Business, and then served as Chair of Department of Management at Bogazici University, Istanbul. His current research interests include management and governance of information systems, IT strategy and business models, open and user innovation, software business and ERP systems. He has published 36 papers in peer-reviewed journals, including *Information Systems Journal*, *Information Economics and Policy*, *Decision Support Systems*, *Empirical Software Engineering*, *Information and Software Technology*, *Electronic Markets*,

Copyright © 2021 Inderscience Enterprises Ltd.

*Information Systems Management*, *Journal of Organizational Computing and Electronic Commerce*, *Journal of Database Management*, *Journal of Software Maintenance and Evolution*, *Enterprise Information Systems*, *Journal of Services Marketing*, *Journal of Global Information Technology Management*, *International Journal of Human Resource Management* and *Wirtschaftsinformatik*, and over 50 in international conference proceedings and book collections.

This paper is a revised and expanded version of a paper entitled 'Multiplicative aggregation for ERP upgrade decision making' presented at International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), Riga, Latvia, 10–12 October 2018.

## 1 Introduction

A problem where one needs to choose one or several best alternatives from a set of many possible alternatives is called a multi-criteria decision making (MCDM) problems. The best alternative is chosen based on several criteria. One decision maker (DM) or a group of persons (DMs) can decide on the choice. We consider a problem where alternatives constitute a discrete set of solutions where each alternative is evaluated using a given set of attributes describing qualities substantial for DMs. The attribute values are known before analysis. The attributes are measures of their utility and are evaluated in definite measurement units. For example, a criterion price can have respective attribute 'price' measured in currency. Criteria define how attributes create the utility of alternatives in multiple dimensions. Thus, attributes create the total utility of the alternative. A criterion catches the utility contribution of an attribute depending on its value. Criteria can be dependent.

MCDM has two distinct fields: multi-attribute decision making (MADM) and multi-objective decision making (MODM) (Pomerol and Barba-Romero, 2000; Triantaphyllou, 2000). Related decision support methods usually express importance of attributes with weights. The problem is expressed in matrix form where each line represents an alternative and columns produce evaluated performance of the attributes.

MADM with particular interest in managerial decision making on IT solutions is considered in the current paper. There is a finite set of alternatives with their criteria. The goal is to rank the alternatives. The ranking is called a complete preorder. Let function $o(\cdot)$ be a precedence aggregation function, and it aggregates criterion utility functions $u_j$ to a single global utility value of DM that is called a unique complete preorder (Pomerol and Barba-Romero, 2000). For $n$ alternatives and $m$ attributes, a general MADM problem can be expressed as the following optimisation problem:

$$\begin{aligned} &\max_i o\left\{u_1(a_{i1}), u_2(a_{i2}), \ldots, u_j(a_{ij}), \ldots, u_m(a_{im})\right\} \\ &\text{such that } g_j(a_{ij}) \le 0,\ j = 1, 2, \ldots\ldots, m, \quad i = 1, 2, \ldots, n \end{aligned} \tag{1}$$

where $i$ and $j$ are indices of alternatives and attributes, $a_{ij}$ are estimated attribute values, $u_j(a_{ij})$ are the criterion objective utility functions, and $g_j(a_{ij})$ are optional model constraints.

An *efficient* alternative has attribute values for respective criteria at least as large as other alternatives have and the value of at least one criterion can be better than the corresponding value of other alternatives. An efficient solution means that any criteria improvement is not possible without reduction of at least one other criteria. A *dominated* solution is a feasible solution, but not an efficient one. The set of all efficient solutions is called the efficient set.

The main goal of this paper is to propose a simple and robust method of managerial decision making assistance for multi-criteria ranking of alternative solutions, and perform preliminary evaluation of its acceptance. We build on a preliminary work by Goman and Koch (2018). Current work finalises development of the composite indicator (CI) and generalises the application of the CI in business MADM. Formulas of the CI are modified in order to become clearer and comprehensible. Additionally, an illustrative example for denominator calculation is supplied, and the example of application of the CI to enterprise resource planning (ERP) upgrade decision was considerably extended to show stability of ranking to scale change of input data and normalisation of weights, and example calculation of suggested formulas is given. The application example is a simplified problem taken from a real-life case study. Empirical evaluation of user acceptance of weighted product (WP) (on which principle the CI is built) in comparison to weighted average (WA) and some formal decision support methods was performed.

The remainder of the paper is organised as follows: the literature review on the problem and related techniques is given in Section 2, the main aggregation approaches (WA and WP) are reviewed in Section 3, the proposed CI based on WP is described in Section 4. Section 5 provides an example of the application of the proposed CI. Results of a first empirical evaluation of user acceptance are given in Section 6. The conclusions summarise the research and provide directions for future research in Section 7.

## 2 Problem background and literature review

MADM uses the function $o(\cdot)$ to aggregate attributes values of alternatives and take into account their weights (relative importance). These aggregates of attribute values can be translated directly into ranks of the alternatives. This aggregation function $o(\cdot)$ is a payoff function of the alternatives. Function $o(\cdot)$ is usually of the form of WA in most decision support methods including most popular ones (e.g., Pomerol and Barba-Romero, 2000; Belton and Stewart, 2002; Tofallis, 2008; Fleming and Wallace, 1986; Qu et al., 2018; Alinezhad and Khalili, 2019). Nature of the criteria and attribute estimates, used scales of data, and aggregation function $o(\cdot)$ determine the meaning of the ranking. Formal decision support methods like analytical hierarchy process (AHP), analytical network process (ANP), ELECTRE, PROMETHEE, TOPSIS, VIKOR, etc. (Triantaphyllou, 2000; Qu et al., 2018; Alinezhad and Khalili, 2019) or simpler procedures based on a CI can be used for ranking of alternatives.

A CI may use more or less formal procedure of alternative comparison that is usually expressed in mathematical formulas. The simplest CI is WA, sometimes also called weighted sum model (WSM). It weights attribute values to produce the aggregated multi-dimensional utility of an alternative. It should be noted that it is presumed that all attributes have the same or directly comparable dimensions, and are subject to additivity assumption.

Another simple CI is WP which is not usually mentioned in literature recently. We refer to a review where it can be found (Kabir et al., 2014) just to imagine the low scale of WP application. However, it has originated in dimensional analysis, and the roots of WP can be traced to the Bridgman's (1931) seminal book in the first part of the 20th century and even earlier. Multiplication of attribute values is used in WP instead of sum in WA, and weights are power orders rather then multipliers in WA.

It should be noted that there is a vast literature on aggregation functions (e.g., Grabisch et al., 2009; Beliakov et al., 2007). Nevertheless, not every aggregation function is directly applicable to MADM in an arbitrary problem domain. In the paper, we do not consider complex problems of weights elicitation and attribute compensation effect. A review on the problems and developments in the area of attribute weighting methods may be found in Pena et al. (2020).

However, the attributes describe different dimensions of the alternative and can be measured in completely different units of measurement. It may happen that the attributes have no common unit of measurement. In order to address that, normalisation procedures are introduced for many decision support methods and CIs like WA. Normalisation is used to ensure that variables with numerically larger magnitude are not dominating, and to remove dimensions of the variables in order that they can be added (to avoid addition of values in different units, which means they are non-comparable). It is used in almost all modern decision support methods (Alinezhad and Khalili, 2019). Sometimes, this operation is called conversion into another (numerical) scale, scaling, standardisation, etc.

There are several normalisation techniques available, and they do not necessarily come to the same result. A normalisation procedure transforms initial attribute vector $a_j$, into a normalised vector $v_j$, $\forall j$, i.e., it assures that all attributes will be in comparable scales. The most known types of normalisation are scaling with maximum value, linear min-max transformation, vector scaling and standardisation ($z$-score), and they can be found in Pomerol and Barba-Romero (2000), Triantaphyllou (2000) and James (2016). Normalisation in general is prone to errors in implementation and result interpretation, which is why our enhancing proposal is based on a method such that the aggregation result does not depend on using any method of normalisation.

Some normalisation procedures (e.g., scaling with maximum value) preserve proportionality. Others (e.g., linear min-max transformation, standardisation) do not preserve proportionality: $a_{ij} / a_{kj}$ is not necessarily equal to $v_{ij} / v_{kj}$, $i = 1, \ldots, n$, $k = 1, \ldots, n$, $i \neq k$. Normalisation also changes the meaning of variables and their impact on the CI. With normalisation, the same CI measured at different times can be non-comparable, so that data should be re-normalised, and the CI recalculated for proper comparison. There are no normalisation rules for ordinal scales and subjective scores (e.g., Hubbard and Seiersen, 2016, Ch. 5). To sum up, normalisation is an artificial step required due to usage of WA aggregation, normalised attribute values do not add any advantage in themselves. To the contrary, there is considerable difficulty in their interpretation, comprehension of their contribution to CI, and bias may be added to ranking because ratios between attribute values may be destroyed.

Decision support methods are often used for evaluation of software investments. Details and reviews can be found in Renkema and Berghout (1997) and Sen et al. (2009). Diverse techniques are used: fuzzy logic or sets (Tzeng and Shen, 2017; Qu et al., 2018; Triantaphyllou, 2000), heuristics, AHP (Pomerol and Barba-Romero, 2000; Triantaphyllou, 2000; Qu et al., 2018), ANP (Qu et al., 2018; Alinezhad and Khalili, 2019), outranking methods (ELECTRE, PROMETHEE) (Pomerol and Barba-Romero, 2000; Qu et al., 2018; Alinezhad and Khalili, 2019), methods based on similarity to an ideal solution (TOPSIS, VIKOR) (Triantaphyllou, 2000; Qu et al., 2018; Alinezhad and Khalili, 2019), hybrid approaches (e.g., Liu et al., 2020), linear, mixed and nonlinear programming, etc. However, most of them use WA directly or indirectly (see method descriptions, e.g., Alinezhad and Khalili, 2019). Unfortunately, normalisation and transformation of initial ratio scale values to ordinal ones is preceding the aggregation for the reasons mentioned above. Moreover, the problems with WA for software selection are already known (e.g., Morisio and Tsoukiàs, 1997).

The rank reversal is discussed and its existence in the most popular decision support methods is indicated with further references in Munier et al. (2019) in Section 2.3. It was shown that those methods that operate on ordinal variables $a_{ij}$ (original data or values after normalisation) or use WA as an aggregation function are known to be prone to rank reversal (Triantaphyllou, 2000; Mahadev et al., 1998; Triantaphyllou and Baig, 2005; Wang and Triantaphyllou, 2006). The best alternative can change after adding or deleting a *non-optimal* alternative. It was also demonstrated that AHP and its modifications are prone to rank reversal [see details in Triantaphyllou, 2000; Hubbard, (2020), p.189]. If TOPSIS is used with non-comparable attributes (having different units of measure), data normalisation is required to compute distance from an ideal alternative. Therefore, it is prone to rank reversal as well. Outranking methods (ELECTRE, PROMETHEE) are complicated, need data normalisation and assume additive aggregation function (which does not have to be always true) (Wang and Triantaphyllou, 2006). ELECTRE (Wang and Triantaphyllou, 2006) and PROMETHEE [Smet, (2019), p.99] are subject to rank reversal.

The diversity of decision support methods and their dependency on transformation of initial attribute values measured in different units of measurement into comparable scales of the criteria values with normalisation procedures can be misleading for practical management decision making: "There are still no rules determining the application of multi-criteria evaluation methods and interpretation of the results obtained" [Zavadskas and Turskis, (2010), p.163]. Furthermore, it is hard to explain operational aspects of fuzzy logic for decision making (e.g., Cooke, 2004), and mathematics of ordinal subjective scores (e.g., used in AHP) is a major challenge (Hubbard and Seiersen, 2016, Ch. 5).

As shown, the choice of a suitable aggregation method to combine the multidimensional set of attributes is an important component in decision support. A managerial decision making requires accuracy of decision support methods and transparency of decision procedures. Misunderstanding value scales and meaning of scores between DMs, confusion in method parameters, and eventual wrong alternative ranking (and imminent improper decision) can be dangerous and expensive. Ratio scales of variables and proper mathematical expressions can greatly minimise the possibility of wrong decision at least by proper formal mathematical constructions and operations, consistently understandable scales of criteria, and unambiguously specified attribute values. A *simple* solution is necessary that can produce *robust and easily explainable* results, and combine data that can be measured in arbitrary scales. In practice, a decision making problem in management has only a *few alternatives*, but their attributes are very different in nature and very often *non-comparable*. Alternatives are known at the time of decision and all relevant attributes are assumed to be known or at least estimated by the time.

## 3 WA and WP

WA of $m$ attributes with weights $w_j \geq 0$ of $n$ alternatives is the most common weighted CI [Triantaphyllou, (2000), pp.6–7]:

$$WA\_CI_i = \sum_{j=1}^{m} w_j a_{ij}, i = 1, \ldots, n; \quad w_j \geq 0, \quad \sum w_j = 1. \tag{2}$$

WA assumes a linear precedence function, independent criteria, weights $w_j$ mean a trade-off ratio between attributes. The latter can be problematic in many practical situations. For example, a DM can assume dependent criteria, e.g., lower price *and* higher usability together with other important factors common to them. It seems unlikely that an alternative with insufficient values of many attributes but with the lowest price attracts much attention in many situations. Naturally, all properties of an alternative are of interest taken together and they are often correlated, i.e., the DM may want higher functionality *and* better return, *and* low cost, *and* high performance, *and* better reliability together. Eventually, weighting does not remove implicit dependence between criteria. Moreover, dependencies in real life are usually nonlinear, so arithmetic aggregation could result in a biased CI.

An aggregation function based on WA is subjected to an effect of criteria substitution. The effect compensates small values of some attributes with high values of other attributes and can produce a competitive result. WA weights therefore constitute a substitution ratio between criteria rather than assume the meaning of importance.

Under an assumption of existence of DM's cardinal function which is additive over the criteria (Pomerol and Barba-Romero, 2000), MADM often handles ordinal evaluations (e.g., 'very bad', 'bad', 'medium', 'good', 'very good') after normalisation that should make the attribute values comparable to other (e.g., numerical) attributes. Nevertheless, application of WA to such attributes means that they *can be additive* in terms of their physical meaning of the application area and *feasibility* of operation of mathematical addition for the normalised scales. The two latter issues are usually not debated for management decision making.

Therefore, WA has deficiencies for application in practical decision making listed in Tofallis (2008). They result is rank reversal, choice of an unbalanced or non-optimal alternative, or reduction of the set of efficient alternatives. Thus, the method is not robust and problematic in practical decision making. Nevertheless, WA can be applied to comparable raw attribute values. Then, post-normalisation to a certain baseline is not problematic and ranking remains the same after linear transformation (Fleming and Wallace, 1986).

For $m$ attributes with weights $w_j$ and $n$ alternatives, a CI based on WP are the following [Triantaphyllou, (2000), pp.8–9]:

$$WP\_CI_i = \prod_{j=1}^{m} a_{ij}^{w_j}, i = 1, \ldots, n. \tag{3}$$

WP can be represented as a weighted sum of logarithms and in general, the sum of the weights need not be equal to unity. WP uses predefined attributes weights but does not require attribute normalisation. Amplification of influence of certain criteria is possible by permitting $w_j > 1$. This means larger influence on CI value of attribute with larger values. To decrease the influence of values from the upper end of the scale relative to ones from the lower end of the scale, one can set $w_j < 1$ (Tofallis, 2008). *WA_CI* is always not smaller than *WP_CI*.

A review of advantages of multiplicative aggregation functions $o(\cdot)$ based on WP is given in Aczél and Alsina (1987). Aggregation with WP is invariant to strictly positive affine mapping [Pomerol and Barba-Romero, (2000), p.53]. Thus, proposed aggregation function (3) is meaningful on independent ratio scales, see Propositions 7.1 and 7.8 in Grabisch et al. (2009, Ch. 7.2).

WP possesses smaller substitutional effect than WA. In this way, an alternative having very low value of at least one criteria will decrease the CI of an alternative sharply irrespective of superior values of other criteria. Contribution of very large values of criteria is also increased. All in all, due to absence of linear compensation for low criteria values, alternatives having very high values of relatively few attributes cannot reach high CI. This should mean better consistency of relative ranking for alternatives with large variance in criteria values. As a result, WP is flexible in adjustment to DM criteria preference and weights. But, most importantly, WP lacks rank reversal due to attribute rescaling when applied properly and this enables comparability across time.

While data normalisation is not needed, it may be applied on condition that it should preserve ratio scale of initial data.

But the most important quality of WP is dimensional soundness, i.e., there is no need to *add* together non-comparable attributes. It aggregates different attributes that have different dimensions correctly, and assures invariance to admissible transformation of scales. For two alternatives having $WP\_CI_1$ and $WP\_CI_2$, it is always possible to consider a dimensionless ratio $WP\_CI_2$ / $WP\_CI_2$. Dimensional soundness is an absolutely necessary *normative principle* for managerial decision making.

## 4 Construction of the CI

The goal is an unbiased and simple aggregation function for criteria of alternatives for decision making. It follows from Section 2 and Section 3 that most known methods based on WA are inappropriate. Attributes are non-comparable, non-additive and can have non-positive values. Therefore, classical WP is not suitable for the task as well. Goman and Koch (2018) suggested a number of conditions for the new CI for strategic decision making, including application of raw data or estimations expressed in ratio scale (absence of data normalisation), established mathematical basis, meaningfulness of aggregation result, easily explainable result of aggregation, and simplicity of the aggregation procedure that should enable ranking of alternatives. Finally, the aggregation method should be robust to permissible scale change.

Ebert and Welsch (2004) and Roberts (1984) demonstrated that a sensible or meaningful CI is not possible for non-comparable ordinal or interval scale values and that a CI based on WA or WP may show rank reversal for these types of variables. This puts contemporary managerial textbooks on decision making into question where normalisation of non-comparable data to ordinal scores or interval scales is performed. However, WP enables strong ordering for both non-comparable and comparable ratio scale values (Ebert and Welsch, 2004). Thus, WP is efficient, for it does not need normalisation and it prevents possible errors due to improper normalisation and operational errors in the normalisation procedure itself. Moreover, verification of ranking is easy with available raw data and known units of measurement. Additionally, WP guarantees scale invariant rankings for ratio scale data.

Majority of attributes in decision making can be of ratio scale type, however, because of different physical sense, they are usually non-comparable. Furthermore, the task of providing a meaningful CI based on WP presumes that such a CI can be obtained only for strictly positive variables $a_{ij} > 0$. Financial results for example can be expressed in monetary terms. Non-financial results usually cannot be expressed in such monetary terms, though sometimes there is a possibility for conversion (e.g., with exchange rates or prices). Positive values would denote benefits; however negative values may exist as well.

Taking into account restrictions of initial data and their type, only ratio type of data is considered hereafter, and the CI is build using WP. Technological, financial, and any other attributes can generally be measured in ratio scale type. To give several examples, one can measure profitability, cash flow, ROI, expenses, project duration, payback period, internal rate of return, effort, performance, scalability, quality of management or support, headcount, data rate, requirements fulfilment, ratios (e.g., expenses over income), etc. in real numbers instead of subjective ordinal scores. Even parameters such as risk (Hubbard and Seiersen, 2016) or return on management (Strassmann, 1999) can be measured using ratio numbers. Let us consider all relevant attributes falling under one of two categories: earnings and expenses. If earnings are less than expenses, the balance is a negative number (i.e., loss). In the opposite, if earnings are greater than expenses, there is profit. Nevertheless, the attribute scales can be non-comparable and the values of different alternatives can differ by several orders of magnitude or be negative.

Should any criterion be a minimisation criterion, its direction can be changed to maximisation by means of the substitution $v_{ij} = 1 / a_{ij}$ that results in $\mathrm{Min}_j(a_{ij}) = \mathrm{Max}_j(v_{ij})$. This is because there are only positive attribute values under consideration. Criteria with undesirable meaning and possibly negative attribute values are treated by modulo of the attribute values $|a_{ij}|$ as in accounting. The reverse is also possible, namely to change a maximisation criteria to a minimisation criterion. This is called *criterion change* operation in the following section. Criterion change operation preserves relational ratio between attribute values [Pomerol and Barba-Romero, (2000), p.53].

Let attributes having maximisation criteria be $S^+$ and attributes that match minimisation criteria be $S^–$. All values that belong to $S^+$ should go into the nominator, for they naturally increase the value of the alternative and therefore, the CI. Attributes belonging to $S^–$ naturally have minimisation criteria, and they are going to the denominator of the CI after application of the criterion change operator, and have maximisation criteria then. Based on that in construction of the CI, we consider transfer from attribute values $a_{ij}$ to new values $b_{ij}$ below with simple functions that do not change their ratio scale properties.

We assume that the scales of attributes in $S^+$ are only positive ratio scales. There should be a good reason to consider alternatives with negative attribute values in $S^+$ as well as with zero values. An alternative with non-existent or null value of an *important* attribute is obviously an exceptional case. However, it can happen in practice, e.g., if all alternatives have unsatisfactory attribute value altogether or certain ones have sufficient quality of all but one attributes. Should virtually non-positive values of attributes $a_{ij} \mid j \in S^+$ occur, a special decision is needed for the situation, e.g., if the value should go to the denominator of the proposed CI for the given alternative assuming criterion change operation or the attribute should be excluded. The transfer from original attribute values $a_{ij} \mid j \in S^{+}$ to new values $b_{ij} \mid j \in S^{+}$ is straightforward (these values are put into the nominator of the CI):

$$b_{ij} = f\left(a_{ij}\right) = \begin{cases} a_{ij} & a_{ij} > 0 \\ error & a_{ij} = 0 \end{cases}. \tag{4}$$

We indicate in the relation above with *error*, the unacceptable case when $a_{ij} = 0$, because the CI based on WP cannot produce a meaningful result with zero argument. Thus, positive values $a_{ij} \mid j \in S^{+}$ are left unchanged.

Attributes that belong to the set $S^{-}$ decrease the CI. The modulo of the values is used similar to *accounting practice*. The function $h(\cdot)$ for attributes that belong to the set $S^{-}$ is the following (these values are put into the denominator of the CI):

$$b_{ij} = h\left(\left|a_{ij}\right|\right) = \begin{cases} \left|a_{ij}\right| \cdot 10^{k} & 0 < \left|a_{ij}\right| < 1 \\ \left|a_{ij}\right| & \left|a_{ij}\right| > 1 \\ error & a_{ij} = 0 \end{cases}. \tag{5}$$

Note, that the physical meaning of attributes in the set $S^{-}$ presumes that their original value should not become positive. For example, there is no reason for expenses to become benefits. If this is so, the respective amount should likely be captured in the nominator. In such a case, additional care should be employed in the modelling stage. This can be a sign of improper criteria choice, and the respective criteria and attributes may need revision. In this situation, an affected attribute, or only a specific alternative may need special treatment within business decision making. For instance, the problematic attribute values can be changed or eliminated with constraints $g(a_{ij})$ [see equation (1)].

Zero values are not allowed in the denominator, therefore this situation is denoted with *error* in relation (5) above. Due to a denominator property, it is suggested to rescale values $0 < |a_{ij}| < 1$ in formula (5) to $k$ number of orders of 10, and $k$ should be chosen such that for a given attribute $j \in S^{-}$, values $b_{ij}$ *for all alternatives* become greater than 1. This is possible, if one chooses such a minimum k that assures that the modulo of the smallest $b_{ij}$ becomes greater than 1: $|a_{ij}| \cdot 10^{k} \geq 1$. For instance, let us assume for a specified attribute $j \in S^{-}$ that $|a_{2j}| = 0.4$ is the second smallest attribute value $\in (0, 1)$, and the smallest attribute value is $|a_{1j}| = 0.01$. The minimum number of orders such that $b_{1j} \geq 1$ and $b_{2j} \geq 1$ is $k = 2$. Therefore, $b_{1\mathrm{j}} = 0.01 \cdot 100 = 1$ and $b_{2j} = 0.4 \cdot 100 = 40$. The transformation preserves the relational ratio between attribute values $b_{1j}$ and $b_{2j}$ the same as between $|a_{1j}|$ and $|a_{2j}|$: $|a_{1j}| \leq |a_{2j}|$, $b_{1j} \leq b_{2j}$, and $|a_{1j}| / |a_{2j}| = b_{1j} / b_{2j}$. It is important to note that if rescaling in formula (5) is applied to any attribute value, the same attribute of other alternatives should be rescaled with the same chosen $k$ as well, as we have just illustrated in the example of two values.

Now, assuming maximisation of all criteria, the expression for the CI for an alternative $i$ is introduced:

$$CI_i = \frac{\prod_{j \in S^{+}} b_{ij}^{w_j}}{\prod_{j \in S^{-}} b_{ij}^{w_j}}; \quad b_{ij} > 0; \quad w_j > 0. \tag{6}$$

In this way, the CI [equation (6)] can process positive ratio scale and non-zero values $b_{ij}$ obtained from initial attributes $a_{ij}$ with formulas (4) and (5), where criteria $j \in S^-$ decrease the CI and $j \in S^+$ increase it. The only requirement is $b_{ij} \geq 1$ for the denominator of the CI, and this is assured in formula (5) for ratio scale data by rescaling $0 < |a_{ij}| < 1 \mid j \in S^-$ with required number of orders $k$. Thus, the new CI can effectively aggregate non-comparable values of different units of measure, and the outcome is meaningful due to independent ratio scales of attributes (Grabisch et al., 2009, Ch. 7.2).

## 5 Application: ERP upgrade decision

For supporting its business processes, every organisation that wants to benefit from an integrated system has to decide on acquisition of standard software or individual solution development. ERP systems are large enterprise information systems with subsystems that enable planning and control of resources and processes, can improve operations and increase competitiveness (Chatzoglou et al., 2016). Standard software solution furthermore presumes selection of the software itself and a vendor for implementation, maintenance and support. This decision will have a major influence on the capabilities of an organisation over a long time period.

Continuous changes in business environment and new emerging technologies put companies under pressure to adapt their business models, strategy, organisational structures, and consequently their information systems. Earlier research has been mainly on the early phases in the life cycle of ERP systems and especially on the implementation (e.g., Hakim and Hakim, 2010). Evolution, maintenance, and replacement of such systems have received comparably less focus, with some exceptions (e.g., Koch and Mitteregger, 2016). The ERP change decision making is an important strategic managerial problem.

Implementation of an ERP system is a challenging process which triggers both technological and organisational changes. A considerable part of the lifecycle is maintenance. ERP upgrade is a regular part of the maintenance process. At times of ERP upgrade there arise similar problems to those at implementation stage (Koch and Mitteregger, 2016). But successful ERP upgrade produces such advantages as better maintenance, defect elimination, improved architecture, costs optimisation, new beneficial features and technologies. At this point, software cannot only be upgraded, but replaced with a competing solution, and other possibilities exist, including change of customisation level, and even postponement of the upgrade (Chatzoglou et al., 2016; Koch and Mitteregger, 2016; Law et al., 2010). The consequence of an inefficient decision can be expensive. ERP upgrade usually includes:

1. identifying main factors that are required for the future system
2. analysing the factors and identifying upgrade options and attributes that describe their properties
3. building a model and analysing factor dependencies
4. aggregating the factors influence into a meaningful value or set of values for ranking the alternatives.

This is a typical MADM problem. The advantage of the CI based on WP over WA will be demonstrated with it. Some especially useful properties are independence from normalisation and ability to work with data in their original scales.

We performed a case study with a large automotive company in their ERP change project. This was a possibility to evaluate the multiplicative CI in real-world managerial ERP decision making. The company is an automotive company that was established more than a century ago where work ca. 3,000 employees, and its average revenue approaches $1 billion per year. The company is one of the largest manufacturers in its specific market segment. There are more than ten production facilities on three continents, and sales and service filial branches in more than 100 countries.

The company ran an in-house developed ERP system that was more than 20 years old. It was basing on a number of piece-wise third party applications, database and system software, and infrastructure solutions. The system was recognised obsolete. The decision to change the system was caused by the anticipation of an inability of its efficient further maintenance and development. Alternative solutions and project implementation partners were re-evaluated several times during ERP change initiative in several phases. The re-evaluation was necessary due to the fact that there are not many partners with expert know-how on all ERP modules, and implementation was required in a number of regions of the world.

Six alternatives for the anonymised ERP upgrade decision problem are defined:

1 A1 upgrade the current ERP system to the newest version (from vendor 1).

2 A2 further customise the current ERP system.

3 A3 replace current ERP system with 1st alternate ERP system (from vendor 2).

4 A4 replace current ERP system with 2nd alternate ERP system (from vendor 3).

5 A5 develop a new custom system.

6 A6 do not change ERP system, but adapt business processes.

Next, ten criteria for the alternatives are defined, which are listed and described in Table 1. Most of the attributes used are measured in non-comparable scales. Most attribute criteria should be minimised and therefore connoting expenses, loss or unfavourable properties. However, their attribute values are nevertheless positive numbers. All alternatives are efficient and there is no dominating solution. One can aggregate the value of each alternative with given criteria weights using the CI. Weights do not add to 1, their sum is 10.4. Weights were normalised such that they add to 1 by division over the sum of weights. It does not change the ranking by CI, but allows to compare CI and WA more directly.

To apply the CI, the criterion change operation is applied to all attributes with minimisation criteria first, so that all criteria be maximisation criteria, and rescaling for attributes whose values were below 1 is applied as described in Section 4. The values used in computations are provided in Table 2. Monetary attributes were entered as thousands of euro. The results of ranking the alternatives are shown in Table 2 with the respective CI values and ranks of each alternative. For illustration, we provide calculation of the CI using formulas (4)–(6) for alternatives A1 and A2 with original weights $w_j$ (from Table 1):

$$CI_1 = \frac{3^2 \cdot 0.5^{0.9} \cdot 500^1}{1000^1 \cdot 700^1 \cdot 100^{0.5} \cdot 150^{0.5} \cdot 12^2 \cdot 2^1 \cdot (2 \cdot 10)^{0.5}} = 21.839,$$
$$CI_2 = \frac{5^2 \cdot 1^{0.9} \cdot 1000^1}{2000^1 \cdot 500^1 \cdot 100^{0.5} \cdot 300^{0.5} \cdot 13^2 \cdot 1^1 \cdot (0.8 \cdot 10)^{0.5}} = 301.959. \tag{7}$$

Due to attribute value 'system reliability' $\in S^-$ is below 1, scaling is required for this attribute *for all alternatives* as specified in equation (5). The smallest $k$ that assures that attribute 'system reliability' for all alternatives becomes not less that 1 is $k$ = 1. Therefore, the respective attribute values for 'system reliability' are multiplied by 10 in formula (7). Finally, all computed CIs and ranks of alternatives are given in Table 2 for both the cases where weights are original and where weights are normed to 1.

**Table 1** Criteria and attributes

| *Factor (abbr.)* | *Units* | *Criteria* | *Weight* | *Comment* |
|---|---|---|---|---|
| Project costs (PC) | € | Desc. | 1 | |
| Support costs (SC) | € | Desc. | 1 | Over 2 years |
| Extra integration expenses (EE) | € | Desc. | 0.5 | Beyond the project – in IT operations |
| Infrastructure risk (IR) | € | Desc. | 0.5 | Estimation of extra purchases |
| Implementation time (IT) | Month | Desc. | 2 | |
| Usability (U) | Second | Desc. | 1 | User time per average operation |
| System reliability (SR) | Rate, 1/year | Desc. | 0.5 | Critical incidents per year |
| Return on investment (ROI) | Dim. less, ratio | Asc. | 2 | |
| Project management practice (PMP) | Dim. less, ratio | Asc. | 0.9 | Rate of successful projects |
| Net present value (NPV) | € | Asc. | 1 | Resulting from changes to business |

The CI values in Table 2 should now reflect the true cumulative utility of each alternative. The dependence is clearly nonlinear: A2 (the best) and A4 (the second best) are favourites, but even they have about double difference in CI. CI produces the same ranking for comparison with original weights and normed weights. Aggregation of the original values with WA is not possible because the attribute values were not normalised to one comparable scale, so they cannot be added.

Results of CI do not depend on the scale of variables as per design. In order to check this, calculations are repeated with all monetary attributes expressed in millions. The result is given in Table 2. CI maintained rank order of alternatives.

In order to illustrate the problems with WA when applied to normalised data, the solutions of the same problem obtained with WA and the proposed CI are compared. The normed weights are used, and attribute values $a_{ij}$ are always recomputed as normalisation methods require. The computation results and alternative ranks are provided in Table 3. CI ranking is the same independently of normalisation procedures, and the same as in Table 2. Normalisation with $z$-score (standardisation) and linear normalisation cannot be applied with WP and the CI due to they produce non-positive values and zeros, respectively. However, we provide WA ranking in Table 3 for linear normalisation of the original data, where WA produces different ranks than before with WA and CI. WA result with linear normalisation is different, i.e. rank reversal occurs.

**Table 2** Attribute estimates and ranks for original data values

| *Factor* | *Unit* | *Weight* | *A1* | *A2* | *A3* | *A4* | *A5* | *A6* |
|---|---|---|---|---|---|---|---|---|
| *Monetary attributes expressed in thousands, original weights $w_j$* | | | | | | | | |
| PC | K€ | 1 | 1,000 | 2,000 | 1,500 | 1,200 | 4,000 | 100 |
| SC | K€ | 1 | 700 | 500 | 200 | 500 | 400 | 100 |
| EE | K€ | 0.5 | 100 | 100 | 200 | 50 | 50 | 10 |
| IR | K€ | 0.5 | 150 | 300 | 200 | 80 | 300 | 20 |
| IT | Month | 2 | 12 | 13 | 18 | 20 | 30 | 1 |
| U | s | 1 | 2 | 1 | 3 | 1.5 | 2 | 10 |
| SR | Ratio | 0.5 | 2 | 0.8 | 0.5 | 1 | 1 | 12 |
| ROI | Ratio | 2 | 3 | 5 | 0.5 | 2.2 | 3 | 0.2 |
| PMP | Ratio | 0.9 | 0.5 | 1 | 0.8 | 0.4 | 0.5 | 0.7 |
| NPV | K€ | 1 | 500 | 1,000 | 1,000 | 5,000 | 10,000 | 50 |
| CI with original $w_j$ (from Table 1), $\cdot\ 10^9$ | | | 21.84 | 301.96 | 1.57 | 147.35 | 43.24 | 93.65 |
| CI rank | | | 5 | 1 | 6 | 2 | 4 | 3 |
| CI, normed $w_j$, $\cdot\ 10^2$ | | | 1.83 | 2.36 | 1.42 | 2.20 | 1.96 | 2.11 |
| CI rank | | | 5 | 1 | 6 | 2 | 4 | 3 |
| *Monetary attributes expressed in millions, normed weights $w_j$* | | | | | | | | |
| PC | Mio € | 0.096 | 1 | 2 | 1.5 | 1.2 | 4 | 0.1 |
| SC | Mio € | 0.096 | 0.7 | 0.5 | 0.2 | 0.5 | 0.4 | 0.1 |
| EE | Mio € | 0.048 | 0.1 | 0.1 | 0.2 | 0.05 | 0.05 | 0.01 |
| IR | Mio € | 0.048 | 0.15 | 0.3 | 0.2 | 0.08 | 0.3 | 0.02 |
| IT | Month | 0.192 | 12 | 13 | 18 | 20 | 30 | 1 |
| U | s | 0.096 | 2 | 1 | 3 | 1.5 | 2 | 10 |
| SR | Ratio | 0.048 | 2 | 0.8 | 0.5 | 1 | 1 | 12 |
| ROI | Ratio | 0.192 | 3 | 5 | 0.5 | 2.2 | 3 | 0.2 |
| PMP | Ratio | 0.086 | 0.5 | 1 | 0.8 | 0.4 | 0.5 | 0.7 |
| NPV | Mio € | 0.096 | 0.5 | 1 | 1 | 5 | 10 | 0.05 |
| CI, normed $w_j$, $\cdot\ 10$ | | | 2.86 | 3.68 | 2.22 | 3.43 | 3.05 | 3.29 |
| CI rank | | | 5 | 1 | 6 | 2 | 4 | 3 |
| CI with original $w_j$ (from Table 1), $\cdot\ 10^6$ | | | 2.18 | 30.20 | 0.16 | 14.74 | 4.32 | 9.37 |
| CI rank | | | 5 | 1 | 6 | 2 | 4 | 3 |

The results in Table 3 show that CI based on WP produces a stable result equal to the ranking in Table 2 in case that normalisation procedure preserves relational ratios in data. Thus, in order to obtain reliable results, a CI based on WP should be used with data normalisation that keeps relative ratio between values. Ranking with WA is stable with data normalisation procedures that keep relational ratios as well, however the ranks can be different than those with CI. With normalisation procedure c (linear normalisation), the WA ranking is very different from the cases with normalisation procedures a and b. But alternative A5 that is visibly inferior in most of the criteria to A2 and A4 is ranked top by the WA. The DM should decide which ranking is correct: obtained with WA or our CI, and which normalisation is justified. Nevertheless, the difference in ranking is evident.

**Table 3** Comparison of WA and CI stability to data normalisation

| *Factor* | *Weight* | *A1* | *A2* | *A3* | *A4* | *A5* | *A6* |
|---|---|---|---|---|---|---|---|
| *a) Normalisation: scaling to the maximum value $v_{ij} = a_{ij} / max_i\, a_{ij}$* | | | | | | | |
| Project costs | 0.096 | 0.250 | 0.500 | 0.375 | 0.300 | 1.000 | 0.025 |
| Support costs | 0.096 | 1.000 | 0.714 | 0.286 | 0.714 | 0.571 | 0.143 |
| Extra expenses | 0.048 | 0.500 | 0.500 | 1.000 | 0.250 | 0.250 | 0.050 |
| Infrastructure risk | 0.048 | 0.500 | 1.000 | 0.667 | 0.267 | 1.000 | 0.067 |
| Implementation time | 0.192 | 0.400 | 0.433 | 0.600 | 0.667 | 1.000 | 0.033 |
| Usability | 0.096 | 0.200 | 0.100 | 0.300 | 0.150 | 0.200 | 1.000 |
| System reliability | 0.048 | 0.167 | 0.067 | 0.042 | 0.083 | 0.083 | 1.000 |
| ROI | 0.192 | 0.600 | 1.000 | 0.100 | 0.440 | 0.600 | 0.040 |
| Project mgmt. practice | 0.086 | 0.500 | 1.000 | 0.800 | 0.400 | 0.500 | 0.700 |
| NPV | 0.096 | 0.050 | 0.100 | 0.100 | 0.500 | 1.000 | 0.005 |
| CI, · 10 | | 1.04 | 1.34 | 0.81 | 1.25 | 1.11 | 1.20 |
| CI rank | | 5 | 1 | 6 | 2 | 4 | 3 |
| WA, · 10 | | 4.36 | 5.74 | 3.88 | 4.36 | 6.82 | 2.41 |
| WA rank | | 4 | 2 | 5 | 3 | 1 | 6 |
| *b) Normalisation: vector scaling* $v_{ij} = \dfrac{a_{ij}}{\left(\sum_i a_{ij}^2\right)^{0.5}}$ | | | | | | | |
| Project costs | 0.096 | 0.201 | 0.402 | 0.302 | 0.241 | 0.805 | 0.020 |
| Support costs | 0.096 | 0.639 | 0.456 | 0.183 | 0.456 | 0.365 | 0.091 |
| Extra expenses | 0.048 | 0.392 | 0.392 | 0.784 | 0.196 | 0.196 | 0.039 |
| Infrastructure risk | 0.048 | 0.300 | 0.601 | 0.401 | 0.160 | 0.601 | 0.040 |
| Implementation time | 0.192 | 0.273 | 0.295 | 0.409 | 0.454 | 0.681 | 0.023 |
| Usability | 0.096 | 0.182 | 0.091 | 0.274 | 0.137 | 0.182 | 0.912 |
| System reliability | 0.048 | 0.163 | 0.065 | 0.041 | 0.081 | 0.081 | 0.977 |
| ROI | 0.192 | 0.432 | 0.721 | 0.072 | 0.317 | 0.432 | 0.029 |
| Project mgmt. practice | 0.086 | 0.299 | 0.599 | 0.479 | 0.239 | 0.299 | 0.419 |
| NPV | 0.096 | 0.044 | 0.089 | 0.089 | 0.443 | 0.886 | 0.004 |
| CI, · $10^2$ | | 7.12 | 9.16 | 5.53 | 8.55 | 7.60 | 8.19 |
| CI rank | | 5 | 1 | 6 | 2 | 4 | 3 |
| WA, · 10 | | 3.05 | 3.98 | 2.74 | 3.13 | 4.98 | 1.96 |
| WA rank | | 4 | 2 | 5 | 3 | 1 | 6 |

**Table 3** Comparison of WA and CI stability to data normalisation (continued)

| *Factor* | *Weight* | *A1* | *A2* | *A3* | *A4* | *A5* | *A6* |
|---|---|---|---|---|---|---|---|
| *c) Linear normalisation* $v_{ij} = \dfrac{a_{ij} - \min_i a_{ij}}{\max_i a_{ij} - \min_i a_{ij}}$ or $v_{ij} = \dfrac{\max_i a_{ij} - a_{ij}}{\max_i a_{ij} - \min_i a_{ij}}$ | | | | | | | |
| Project costs | 0.096 | 0.769 | 0.513 | 0.641 | 0.718 | 0.000 | 1.000 |
| Support costs | 0.096 | 0.000 | 0.333 | 0.833 | 0.333 | 0.500 | 1.000 |
| Extra expenses | 0.048 | 0.526 | 0.526 | 0.000 | 0.789 | 0.789 | 1.000 |
| Infrastructure risk | 0.048 | 0.536 | 0.000 | 0.357 | 0.786 | 0.000 | 1.000 |
| Implementation time | 0.192 | 0.621 | 0.586 | 0.414 | 0.345 | 0.000 | 1.000 |
| Usability | 0.096 | 0.889 | 1.000 | 0.778 | 0.944 | 0.889 | 0.000 |
| System reliability | 0.048 | 0.870 | 0.974 | 1.000 | 0.957 | 0.957 | 0.000 |
| ROI | 0.192 | 0.583 | 1.000 | 0.063 | 0.417 | 0.583 | 0.000 |
| Project mgmt. practice | 0.086 | 0.167 | 1.000 | 0.667 | 0.000 | 0.167 | 0.500 |
| NPV | 0.096 | 0.045 | 0.095 | 0.095 | 0.497 | 1.000 | 0.000 |
| WA, · 10 | | 5.03 | 6.50 | 4.40 | 5.08 | 4.40 | 5.24 |
| WA rank | | 4 | 1 | 5 | 3 | 6 | 2 |

There is no point in illustrating WA application to data with all attribute values normalised to respective values of one of the alternatives as in Fleming and Wallace (1986). Normalisation procedures were explained in Fleming and Wallace (1986), and this is a division of values of the attribute of all alternatives over the attribute value of the alternative chosen as the baseline. For instance, we can norm all respective attribute values to alternative A1: 'project costs' of all alternatives is divided by the value of 'project costs' of alternative A1, 'usability' is divided by the value of the 'usability' attribute of alternative A1, 'NPV' of each alternative is divided by the value of the respective attribute of alternative A1, etc. WA should not be applied after such normalisation, because normalisation cannot bring all attributes to the same scale (they are still non-comparable). In fact, data normalised in this way are of different orders, and still have diverse physical meaning. The difference between the example from Fleming and Wallace (1986) and the one considered here is that now this *is a multi-criteria* comparison with *many non-comparable attributes*. WA cannot be applied to add NPV, usability, ROI or something else directly. There is little sense in WA application to such data, and there exists inevitable rank reversal with WA. Nevertheless, the CI can be applied to data normalised in this way, and CI ranking is *absolutely* the same as in other cases shown before.

## 6 Empirical evaluation of user acceptance

In presence of multiple approaches to aggregate non-comparable attributes to determine value of several alternatives that can return various ranking of alternatives (and may depend on data normalisation), we consider that one should ask those users (DMs) who will use the methods about their preference. For evaluation of user acceptance, a survey was chosen to understand the relation between comprehension and willingness to use of

WA, WP and a number of other formal decision support methods, and to compare them accordingly. They are summarised in Table 4. The objective was to evaluate how well people understand decision support methods and on that basis express their willingness to apply them in practice. Thirty-three students were interviewed who are currently enrolled in a class on IT project management as part of a business informatics bachelor program. All the students have studied a number of managerial subjects and have basic knowledge about ERP and decision making approaches. A number of students also have practical experience at companies. All of the respondents knew about WA method for ranking, 42% were aware of WP, and 21% of classical AHP. All answers were anonymous.

**Table 4** Decision support methods

| *Method* | *Reference* |
|---|---|
| Classical AHP (AHP1) | Triantaphyllou (2000), Qu et al. (2018) |
| Revised AHP (AHP2) | Triantaphyllou (2000) |
| ELECTRE | Triantaphyllou (2000), Pomerol and Barba-Romero (2000), Qu et al. (2018), Alinezhad and Khalili (2019) |
| PROMETHEE | Pomerol and Barba-Romero (2000), Qu et al. (2018), Alinezhad and Khalili (2019) |
| VIKOR | Qu et al. (2018), Alinezhad and Khalili (2019) |
| TOPSIS | Triantaphyllou (2000), Pomerol and Barba-Romero (2000), Qu et al. (2018) |

It was not possible to identify any previous works on comparable evaluation of decision method acceptance. The closest areas seem to be satisfaction measurement from Bernroider and Schmöllerl (2013), and usability measurement from Tullis and Albert (2013). The interest lies in ex-ante expectations of users regarding different methods based on information about them acquired within a questionnaire. Thus, the questionnaire was developed according to the concepts from both of the sources above as follows: after a short abstract about MADM problems, a simple decision problem from Triantaphyllou and Mann (1989) with description of attributes was provided. There are three alternatives and three attributes in the problem. It is assumed that all the attributes meaning different costs in €. Therefore, the attributes are comparable. The respondents were asked to rank the alternatives themselves based on their subjective opinion. Then, an extensive description of every method was given. The description included relevant textual information about stages, algorithms, formulas, etc. A table with final ranks for alternatives of the sample problem obtained with the method was provided at the end of each method's section. Intermediary calculations were given for complex methods. The following questions followed each method using five-level Likert ordinal scale [yes, rather yes (RY), not exactly (NE), rather no (RN) and no]:

- Do you think that you understand the logic of the method?
- Do you agree with the final ranking?
- Has the method revealed new useful information for decision making that strongly supports the consistency of ranking?
- Do you agree that the method improved your understanding of which alternatives are better than others?

- Do you believe that the method is reliable?

Finally, one question with binary scale followed each method description:

- Are you willing to use the method?

The respondent's answers to the first (understanding the logic of the method) and the last (willingness to use the method) questions are summarised in Table 5. The column 'positive' for each method comprises the sum of 'yes' and 'RY' answers. The column 'WTU' contains the number of affirmative answers to willingness to use question.

**Table 5** Summary of answers about method understanding and willingness to use

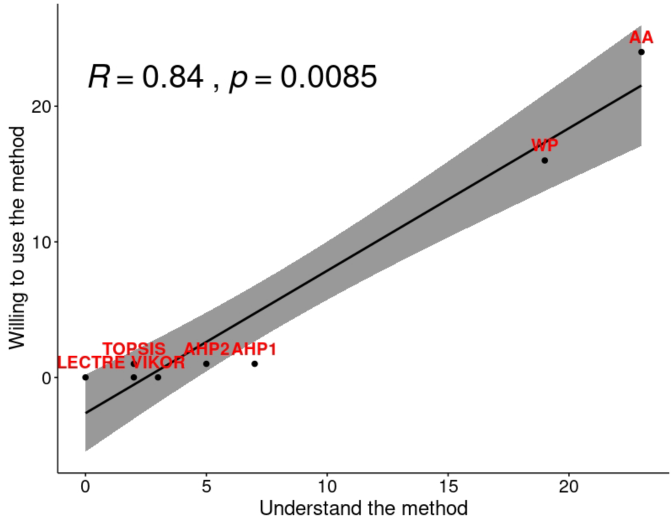
| *Method* | *Yes* | *RY* | *NE* | *RN* | *No* | *Positive* | *WTU* |
|---|---|---|---|---|---|---|---|
| WA | 15 | 8 | 4 | 3 | 3 | 23 | 23 |
| WP | 3 | 16 | 4 | 5 | 5 | 19 | 16 |
| Classical AHP | 1 | 6 | 3 | 0 | 1 | 7 | 1 |
| Revised AHP | 2 | 3 | 3 | 1 | 2 | 5 | 1 |
| ELECTRE | 0 | 0 | 1 | 3 | 7 | 0 | 0 |
| PROMETHEE | 0 | 2 | 5 | 1 | 3 | 2 | 0 |
| VIKOR | 0 | 3 | 3 | 1 | 4 | 3 | 0 |
| TOPSIS | 0 | 2 | 4 | 1 | 4 | 2 | 1 |
| Runs test, $p = 0.05$ | 0.45 | 1 | 0.45 | 0.32 | 0.45 | 0.45 | 0.51 |
| $\chi^2$-test, $\chi^2_{7,0.05} = 14.07$ | 70.05 | 36.40 | 2.93 | 10.07 | 6.59 | 67.66 | 108.07 |

The values of Likert ordinal scales of each question are considered independent variables of proportion of respondents that have chosen the answer to the question about each method. Thus, the methods are evaluated as multidimensional variables in respect to discrete dimensions of user's opinion. The respective dependent variables for the methods are answers to the question about willingness to use each method. The existence of monotonic dependency was checked between concentration of positive answers on the Likert scale and the dependent variable (willingness to use).

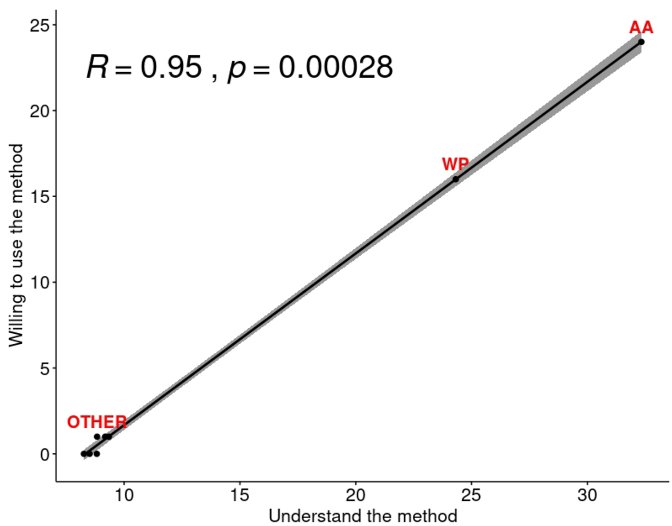
First, the case is evaluated where people express that they understand the method and then answer that they wish to use it. The following answers to the question of understanding are considered positive answers: 'yes' and 'RY', so it is assumed that their sum is an independent variable. Shapiro-Wilk normality test reveals that the values do not likely originate from normal distribution. Therefore, Spearman's rank correlation test was performed on the multidimensional variables in RStudio (version 1.1.419). The correlation between number of positive answers to the question of understanding and willingness to use the method is high: $\rho = 0.8435$, p-value = 0.00848.

Next, assuming linear dependency between all dimensions of answers and the resulting number of people who expressed their wish to use the respective method, multiple regression is employed between the independent variables (proportion of respondents who chose each answer type for a certain decision support method) and the dependent variable (willingness to use the method). A linear model is built and the Spearman correlation evaluated for it. The multiple correlation turned out to be high: $\rho = 0.9511$, p-value = 0.00028.

**Figure 1** Regression between number of people who understand the method and is willing to use it, (a) regression between positive answers and willingness to use (b) multiple regression assuming all types of answers (see online version for colours)




(a)




(b)

Visual presentation of the correlation is provided in Figure 1. It is evident that understanding of the method and willingness to use it are dependent in the experiment. Clearly, WA and WP are dominating other more complex and cumbersome methods. Moreover, there is a gap between the methods perceived as simple (namely, WA and WP) and all other methods on the basis of understanding and the following expression of willingness to use them in practice. Taking into consideration, inconsistencies of many formal decision support methods, problems concealed in normalisation from Section 2, and the advantages of WP (Section 3) for most decision making applications, one can conclude that these two simple aggregation methods turn out to be the most reliable and useful in practice (while WP has not been well promoted).

A null hypothesis $H_0$ is that people's understanding of methods does not make a difference for whether they are willing to use them in practical application. The alternative hypothesis $H_1$ is that they prefer methods that they understand and therefore are willing to choose them for practical application.

The results of a rank sign test (see Table 5) show that the answers (data in columns) are not a random fluke: Wald-Wolfowitz runs test shows that all data are not a random pattern (p-value > 0.05). Provided there is very good correlation between columns, chi-square test for compatibility of *K* counts in Table 5 reveals that answers about understandability and willingness to use are all significantly different for all the methods. Moreover, both the understanding of the methods is significantly different and the respective answers of willingness to use are as well. Therefore, the null hypothesis is rejected and it is believed that the alternative hypothesis $H_1$ is true: people prefer methods that they understand and therefore are willing to choose them for practical application.

Furthermore, two-proportion chi-square test demonstrates that understanding of WA and WP are not significantly different (p-value > 0.05) from each other, whereas all other methods are significantly different from them (p-value < 0.05). The same is true for the test for willingness to use. Computations were done in RStudio (version 1.1.419).

It is interesting to note that chi-square test in Table 5 for positive answers ('yes', 'RY' alone and 'positive') show that people understand all methods statistically differently. However, if they do not understand the methods, the distribution of their answers is the same ('NE', 'RN' and 'no').

## 7 Conclusions

In this paper, a new approach for MADM problems in business decision making is proposed. Additive aggregation is problematic if attributes have different dimensions and scales. In this case, normalisation of values is necessary, which could violate proportionality between attribute values and makes interpretation of the resulting values complicated. Therefore, the proposed method for managerial decision making works with non-comparable attributes, and aggregation based on multiplicative principle seems a candidate which is also robust to rank reversal.

A CI based on WP is introduced. The CI is a mathematically correct way of dimensional aggregation of non-comparable attributes for decision making, and does not cause rank reversal caused by trivial change in attribute scales. The CI operates with maximisation and minimisation criteria, which are in the nominator, respectively the denominator. The only requirement is that values in the denominator must not be below 1, which is assured through rescaling. Moreover, it simplifies the analysis because it needs no normalisation techniques for raw data. Furthermore, verification of the ranking is simplified due to availability of the raw attribute data in the formulas.

If weights add to unity, any of the attributes has less than linear effect on CI growth. Larger attribute values and higher criteria weights produce higher CI value. Attributes with substantially smaller values contribute much less to the overall CI value. The influence is therefore nonlinear. Weights bigger than unity can increase this effect even more because the contribution is now a power function. This approach therefore is flexible and allows CI adjustment for particular problem conditions.

We illustrated method application with a simplified real-life case of ERP upgrade decision making having ten criteria and six alternatives. The stakeholders recognised the advantages of the proposed CI, including ease of application, and acknowledged its superiority over their current methodology and decision support methods they knew. Nevertheless, this is only the first case study in the area of managerial decision making. More empirical studies are required to understand the generality of applicability in decision making problems and user acceptance, e.g., willingness to replace the current practice with the use of the proposed CI in their everyday work.

The experimental comparison between traditionally recommended WA-based CI for decision making and our WP-based CI with an example ERP upgrade problem was performed. It proved theoretical notions of robustness of the WP-based CI to allowable changes of attribute scales. The new CI persuasively evidences stable and transparent outcome independently of permissible scale change of the input data. Computational procedure is simpler than in most of decision support methods, and it needs not normalisation of data.

Empirical evaluation of user acceptance showed that there exists a dependency between comprehensibility of a decision support method and willingness to use it. Simple methods surpass complex formal ones in willingness to use even after the users are familiarised with them. The consequence for decision making is the need of WP promotion as the only available tool to produce comprehensive and robust CI also offering sufficient understandability. However, our study is of small-scale and was performed not among the target group – managers. This result encourages further user acceptance experiments. We believe that similar studies with practitioners in management decision making in diverse business fields are required.

In further research, user acceptance of the CI should be further evaluated in an empirical manner, presumably, in real-life problems in companies, and traditional methods of decision support should participate in the comparison. This can include its application in a variety of current real-life problems or retrospective re-evaluation of already completed decisions with known related information that DMs used. There is a need to compare to the results of other theoretical decision support methods proposed in the literature, including DM's personal attitude to the aspects of correctness, points of comparison and reasoning about usefulness.

Another direction for future work is further detailed evaluation of people's acceptance of decision support methods and willingness to use them as well as ease of results interpretation and adoption of the resulting ranking. As people are DMs, only they can distinguish between a correct and a biased or inconclusive ranking. There is a potential to reveal subjective criteria of acceptability of decision support techniques and correctness of ranking of alternatives for managers. It seems that this is a multidisciplinary area, not purely a technological or mathematical one.

## Acknowledgements

This research was supported by the Austrian Research Promotion Agency (FFG) together with ReqPOOL GmbH under Project Number 866310.

## References


Aczél, J. and Alsina, C. (1987) 'Synthesizing judgements: a functional equations approach', *Mathematical Modelling*, Vol. 9, Nos. 3–5, pp.311–320 [online] http://linkinghub.elsevier.com/retrieve/pii/0270025587904878.

Alinezhad, A. and Khalili, J. (2019) *New Methods and Applications in Multiple Attribute Decision Making (MADM), Vol. 277 of International Series in Operations Research & Management Science*, Springer International Publishing, Cham, Switzerland [online] http://link.springer.com/10.1007/978-3-030-15009-9.

Beliakov, G., Pradera, A. and Calvo, T. (2007) *Aggregation Functions: A Guide for Practitioners, Vol. 221 of Studies in Fuzziness and Soft Computing*, Springer, Berlin, Heidelberg [online] https://doi.org/10.1007/978-3-540-73721-6.

Belton, V. and Stewart, T.J. (2002) *Multiple Criteria Decision Analysis – An Integrated Approach*, Springer US, Boston, MA.

Bernroider, E. and Schmöllerl, P. (2013) 'A technological, organisational, and environmental analysis of decision making methodologies and satisfaction in the context of IT induced business transformations', *European Journal of Operational Research*, Vol. 224, No. 1, pp.141–153 [online] https://doi.org/10.1016/j.ejor.2012.07.025.

Bridgman, P.W. (1931) *Dimensional Analysis*, Yale University Press, New Haven, Connecticut.

Chatzoglou, P.D., Chatzoudes, D., Fragidis, L.L. and Symeonidis, S. (2016) 'Critical success factors for ERP implementation in SMEs', in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, FedCSIS 2016*, Gdańsk, Poland, 11–14 September, pp.1243–1252.

Cooke, R.M. (2004) 'The anatomy of the squizzel: the role of operational definitions in representing uncertainty', *Rel. Eng. & Sys. Safety*, Vol. 85, Nos. 1–3, pp.313–319 [online] https://doi.org/10.1016/j.ress.2004.03.019.

Ebert, U. and Welsch, H. (2004) 'Meaningful environmental indices: a social choice approach', *Journal of Environmental Economics and Management*, Vol. 47, No. 2, pp.270–283 [online] https://linkinghub.elsevier.com/retrieve/pii/S0095069603001177.

Fleming, P.J. and Wallace, J.J. (1986) 'How not to lie with statistics: the correct way to summarize benchmark results', *Communications of the ACM*, Vol. 29, No. 3, pp.218–221 [online] http://portal.acm.org/citation.cfm?doid=5666.5673.

Goman, M. and Koch, S. (2018) 'Multiplicative aggregation for ERP upgrade decision making', in *2018 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, pp.1–7.

Grabisch, M., Marichal, J-L., Mesiar, R. and Pap, E. (2009) *Aggregation Functions, Encyclopedia of Mathematics and its Applications*, Cambridge University Press, New York, New York.

Hakim, A. and Hakim, H. (2010) 'A practical model on controlling the ERP implementation risks', *Inf. Syst.*, Vol. 35, No. 2, pp.204–214 [online] https://doi.org/10.1016/j.is.2009.06.002.

Hubbard, D.W. (2020) *The Failure of Risk Management: Why it's Broken and How to Fix it*, 2nd ed., Wiley, Hoboken, New Jersey.

Hubbard, D.W. and Seiersen, R. (2016) *How to Measure Anything in Cybersecurity Risk*, John Wiley & Sons, Hoboken, New Jersey.

James, S. (2016) *An Introduction to Data Analysis Using Aggregation Functions in R*, Springer International Publishing, Cham, Switzerland [online] http://link.springer.com/10.1007/978-3-319-46762-7.

Kabir, G., Sadiq, R. and Tesfamariam, S. (2014) 'A review of multi-criteria decision-making methods for infrastructure management', *Structure and Infrastructure Engineering*, Vol. 10, No. 9, pp.1176–1210.

Koch, S. and Mitteregger, K. (2016) 'Linking customisation of ERP systems to support effort: an empirical study', *Enterprise Information Systems*, Vol. 10, No. 1, pp.81–107 [online] http://www.tandfonline.com/doi/full/10.1080/17517575.2014.917705.

Law, C.C.H., Chen, C.C. and Wu, B.J.P. (2010) 'Managing the full ERP life-cycle: considerations of maintenance and support requirements and IT governance practice as integral elements of the formula for successful ERP adoption', *Computers in Industry*, Vol. 61, No. 3, pp.297–308 [online] https://doi.org/10.1016/j.compind.2009.10.004.

Liu, Y., Eckert, C.M. and Earl, C. (2020) 'A review of fuzzy AHP methods for decision-making with subjective judgements', *Expert Systems with Applications*, Vol. 161, p.113738.

Mahadev, N.V.R., Pekec, A. and Roberts, F.S. (1998) 'On the meaningfulness of optimal solutions to scheduling problems: can an optimal solution be nonoptimal?', *Operations Research*, Vol. 46, No. 3, Supplement 3, pp.S120–S134 [online] https://doi.org/10.1287/opre.46.3.S120.

Morisio, M. and Tsoukiàs, A. (1997) 'Iusware: a methodology for the evaluation and selection of software products', *Software Engineering – IEE Proceedings*, Vol. 144, No. 3, pp.162–174 [online] https://doi.org/10.1049/ip-sen:19971350.

Munier, N., Hontoria, E. and Jiménez-Sáez, F. (2019) *Strategic Approach in Multi-criteria Decision Making: A Practical Guide for Complex Scenarios, Vol. 275 of International Series in Operations Research & Management Science*, Springer International Publishing, Cham, Switzerland [online] http://link.springer.com/10.1007/978-3-030-02726-1.

Pena, J., Nápoles, G. and Salgueiro, Y. (2020) 'Explicit methods for attribute weighting in multi-attribute decision-making: a review study', *Artificial Intelligence Review*, Vol. 53, No. 5, pp.3127–3152 [online] https://doi.org/10.1007/s10462-019-09757-w.

Pomerol, J-C. and Barba-Romero, S. (2000) *Multicriterion Decision in Management Principles and Practice*, Springer US, Boston, MA.

Qu, Z., Wan, C., Yang, Z. and Lee, P.T-W. (2018) 'A discourse of multi-criteria decision making (MCDM) approaches', in Lee, P.T-W. and Yang, Z. (Eds.): *Multi-criteria Decision Making in Maritime Studies and Logistics: Applications and Cases, Vol. 260 of International Series in Operations Research & Management Science*, pp.7–29, Springer International Publishing [online] https://doi.org/10.1007/978-3-319-62338-2_2.

Renkema, T.J. and Berghout, E.W. (1997) 'Methodologies for information systems investment evaluation at the proposal stage: a comparative review', *Information and Software Technology*, Vol. 39, No. 1, pp.1–13 [online] http://linkinghub.elsevier.com/retrieve/pii/0950584996850063.

Roberts, F. (1984) 'On the theory of meaningfulness of ordinal comparisons in measurement', *Measurement*, Vol. 2, No. 1, pp.35–38 [online] https://linkinghub.elsevier.com/retrieve/pii/0263224184900216.

Sen, C.G., Baraçli, H. and Sen, S. (2009) 'A literature review and classification of enterprise software selection approaches', *International Journal of Information Technology and Decision Making*, Vol. 8, No. 2, pp.217–238 [online] https://doi.org/10.1142/S0219622009003351.

Smet, Y.D. (2019) 'Beyond multicriteria ranking problems: the case of PROMETHEE', in Doumpos, M., Figueira, J.R., Greco, S. and Zopounidis, C. (Eds.): *New Perspectives in Multiple Criteria Decision Making: Innovative Applications and Case Studies, Multiple Criteria Decision Making*, pp.95–114, Springer International Publishing, Cham, Switzerland [online] http://link.springer.com/10.1007/978-3-030-11482-4_3.

Strassmann, P. (1999) *Information Productivity: Assessing Information Management Costs of U. S. Corporations*, Information Economics Press, New Canaan.

Tofallis, C. (2008) 'Selecting the best statistical distribution using multiple criteria', *Computers & Industrial Engineering*, Vol. 54, No. 3, pp.690–694 [online] https://linkinghub.elsevier.com/retrieve/pii/S0360835207001891.

Triantaphyllou, E. (2000) *Multi-criteria Decision Making Methods: A Comparative Study, Vol. 44 of Applied Optimization*, Springer US, Boston, MA.

Triantaphyllou, E. and Baig, K. (2005) 'The impact of aggregating benefit and cost criteria in four MCDA methods', *IEEE Transactions on Engineering Management*, Vol. 52, No. 2, pp.213–226.

Triantaphyllou, E. and Mann, S.H. (1989) 'An examination of the effectiveness of multidimensional decision-making methods: a decision-making paradox', *Decision Support Systems*, Vol. 5, No. 3, pp.303–312.

Tullis, T. and Albert, W. (2013) *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, Morgan Kaufmann Series in Interactive Technologies*, 2nd ed., Morgan Kaufmann, Amsterdam [online] http://www.sciencedirect.com/science/book/9780124157811.

Tzeng, G-H. and Shen, K-Y. (2017) *New Concepts and Trends of Hybrid Multiple Criteria Decision Making*, 1st ed., CRC Press, Boca Raton, Florida [online] https://www.taylorfrancis.com/books/9781351680639.

Wang, X. and Triantaphyllou, E. (2006) 'Ranking irregularities when evaluating alternatives by using some ELECTRE methods', *Omega*, Vol. 36, No. 1, pp.45–63.

Zavadskas, E. and Turskis, Z. (2010) 'A new additive ratio assessment (ARAS) method in multicriteria decision-making', *Technological and Economic Development of Economy*, Vol. 16, No. 2, pp.159–172.